2008 IEEE/RSJ International Conference on Intelligent Robots and Systems
Acropolis Convention Center
Nice, France, Sept, 22-26, 2008

# Laban Movement Analysis for Multi-Ocular Systems

Joerg Rett and Luís Santos and Jorge Dias

***Abstract*— We present as a contribution to the field of human-machine interaction a system that analyzes human movements online through multiple observers, based on the concept of Laban Movement Analysis (LMA). The implementation uses a Bayesian model for learning and classification, while the results are presented for the application to analyze expressive movements. In sports like Karate four judges are placed in the corners to observe the fight to ensure that the overall judgment is correct. In this paper we propose a multi-ocular system where each sub-system observes a movement from a different monocular perspective. The sub-systems send continuously guesses in form of probability distributions to the central system. The central system fuses the evidences and presents the final result. We present the Laban Movement Analysis as a concept to identify useful features of human movements to classify human actions. The movements are extracted using both, vision and magnetic tracker. The descriptor opens possibilities towards expressiveness and emotional content. To solve the problem of classification we use the Bayesian framework as it offers an intuitive approach to learning and classification. The presented work targets applications like social robots, smart houses and surveillance.**

## I. Introduction

The research field of computational Human Movement Analysis is lacking a general underlying modeling language [1]. A semantic descriptor allows to pose the classification task as a problem to recognize a sequence of symbols taken from an alphabet consisting of motion-entities. Systems which are based on such a modeling language can use it as a ground truth for recoding and labeling training data. The inherent constraints of a modeling language can be used to make the task of movement recognition more tractable.

The framework of Laban Movement Analysis (LMA) has already been suggested as a semantic descriptor for human-robot interaction in [2]. Their system implemented the *Space* component of LMA through the concept of *Vector Symbols*. The low-level features were extracted from a monocular camera mounted on the mobile robot 'Nicole'. For tracking the color-based algorithm CAMshift [3] was used. Their classifier was based on a Bayesian approach which allowed a 'online' recognition of six gestures. The robustness of the probabilistic algorithm allowed a trial in a natural environment with untrained actors. The work also introduced *anticipation* as a characteristic of the system emerging from the probabilistic algorithm.

This work is partially supported by FCT-Fundação para a Ciência e a Tecnologia Grant #12956/2003 to J. Rett and by the BACS-project-6th Framework Programme of the European Commission contract number: FP6-IST-027140, Action line: Cognitive Systems to J.Rett and L. Santos

J. Rett and L. Santos and J. Dias are with the Institute of Systems and Robotics, Department of Electrical Engeneering, Pólo II, 3000 Coimbra, Portugal (jrett, luis, jorge)@isr.uc.pt

The aim of this article is to show that the concept of movement classification based on LMA [2] can be applied to multi-ocular systems. We will follow the Bayesian approach to benefit from its robustness for the tasks of multi-sensor data integration and movement capture based on vision. For the former we will follow a hybrid approach by i) recording labeled 3-D movement data using a commercial motion capture device, ii) mapping the data to 2-D planes aligned with the camera planes, iii) learning the LMA descriptors of these planes, iv) classifying 'online' an unknown movement by each of the cameras and finally v) fusing the (pre-)results of the observer systems by taking into account the 'certainty' of their beliefs.

Our system benefits from the simpler processing and higher precision of the commercial motion capture device ('active sensor' [1]) during the recoding step, while using the attractive touch-free alternative of computer vision for the classification step.

The presented system does not require an accurate camera calibration, specially the effect of lens distortions has been neglected. This robustness is partially due to the type of descriptors (*Vector Symbols*) and partially due to the probabilistic approach. The probabilistic approach adds to this robustness by learning features from several trials and persons. This produces probabilities of a certain amount for neighboring values. Additionally, a less accurate calibration allows us to pose less demands on the accuracy of the visual tracking.

The problems that arise when movements are observed by vision, e.g. occlusions, are solved through the Bayesian sensor fusion. A camera which detects the disappearance of a tracked object (e.g. through occlusion) will continue its update with uniformly distributed descriptors resulting in an increasing uncertainty. The algorithm for sensor fusion will assign a lower confidence value having the effect that observers without occlusions are weighted stronger.

In [4] 3-D data from an active sensor was used to obtain a set of movement sequences. Then 2-D projections from several orientations are generated. For the same orientations, projections of a 3-D model to images are created. Treating the two sets as input-output data a neural network was trained. The problem of data association is not trivial due to ambiguity and was solved by clustering statistically homogeneous data points in the 2-D projected marker space. They achieved good results for training five sequences sampled at 32 orientations. As their system only provides the pose of a human body, the classification of movements still remained a open issue and consequently no descriptor was introduced. Also, online or real-time behavior was not addressed in their work.

Section II presents the geometric model and the concept of calibration. Section III presents the concepts of LMA with focus on the *Space* component and the database of expressive movements. Section IV shows the design of the model, learning and classification in the framework of an Bayesian approach. Section VI closes with a discussion and an outlook for future works.

## II. Geometric Model

### A. Projective Matrix

The projective matrix, is a matrix that represents the relationship between two different spaces. The pinhole camera model [5] establishes the geometric relationship between a point in a 3-D object space and it's representation in a 2-D image space. The relationship between image space $I = (u, v, 1)$ and object space $O = (x, y, z, 1)$ in homogeneous coordinates is given by a projective matrix H as follows

$$I^T = HO^T$$

The method used to compute this matrix was Direct Linear Transform (DLT) [5]. The mathematical details of this methods are out of the scope of this work, hence will not be discussed. However this is a data processing tool that will provide an effective way to map points from an Object Space into an Image Space. This model doesn't deal with lens distortion, but that fact was considered not to be relevant considering our human movement model, based in spatial displacements (in form of euclidean distances). As it will be seen in the next chapter, the displacements are direction vectors discretized by 45 degrees, thus exact positions do not pose a real problem to our model, and the approximation taken is considered valid to some extent.

### B. Using Calibration to learn rotated Files

The gesture recognition algorithm is divided in three principal steps: i) Acquisition of movement data, ii) Learning of specific 2-D trajectories and iii) Classification of movements from 2-D camera data. For the first step movement data is acquired by using an active sensor. Three sensors of are attached to the hands ($lh$, $rh$) and the face ($f$) as shown in Fig. 1 a). The 3-D trajectories of several persons and trials are stored for further use. In the Learning step the data is projected to 2-D planes e.g. the principal planes *Door Plane* (vertical) $\pi_v$, *Table plane* (horizontal) $\pi_h$, and the *Wheel Plane* (sagittal) $\pi_s$ as shown in Fig. 1. Working with a static configuration like the principal planes poses a problem for the Recognition step in situations where the real orientation of the cameras is not aligned with the planes.

Another approach was made in order to make the Learning stage as robust as possible preventing such phenomenon from happening. Considering that the cameras used in the Classification step are calibrated and remain static relative to a base referential, the trajectory was learned from the camera's point of view, i.e. the trajectory is mapped into each camera, as it can be seen in Figure 1 b), there is 3-D trajectory in a base referential and its correspondent 2-D mapped trajectory as seen from both cameras 1 and 2. This means that in the Classification stage, matching will occur based on parallel planes.

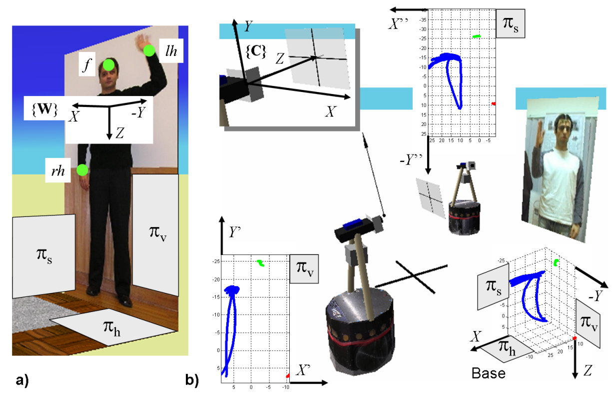


Fig. 1. a) Data acquisition. Three active sensors (green dots) attached to the hands ($lh$ and $rh$) and the face ($f$). Orientation of the three principal planes $\pi_v$, $\pi_h$ and $\pi_s$. b) Classification. Two cameras tracking $lh$, $rh$ and $f$ in 2-D aligned with $\pi_v$ and $\pi_h$.

The presented approach opens the possibility to pose the problem of occlusion as a problem of perspective: An occlusion that occurs in one camera (perspective) might not occur in the other. Furthermore the system can speed up the Classification as the learning of 2-D descriptors from 3-D data can be done offline.

## III. Laban Movement Analysis

Laban Movement Analysis (LMA) is a method for observing, describing, notating, and interpreting human movement. It was developed by a German named Rudolf Laban (1879 to 1958), who is widely regarded as a pioneer of European modern dance and theorist of movement education [6]. The general framework was described in 1980 by Irmgard Bartenieff a scholar of Rudolf Laban in [7]. A computational model of gesture acquisition and synthesis to learn motion qualities from live performance has been proposed in [8]. In neuroscience the usefulness of LMA to describe certain effects on the movements of animals and humans [9] has been inverstigated. The theory of LMA consists of several major components. The works of Norman Badler's group [6] mention five major components. *Relationship* describes modes of interaction with oneself, others, and the environment (e.g. facings, contact, and group forms). *Body* specifies which body parts are moving, their relation to the body center , the kinematics involved and the emerging locomotion. *Space* treats the spatial extent of the mover's *Kinesphere* (often interpreted as reach-space) and what form is being revealed by the spatial pathways of the movement. *Effort* deals with the dynamic qualities of the movement and the inner attitude towards using energy. *Shape* is emerging from the *Body* and *Space* components and focused on the body itself or directed towards a goal in space.

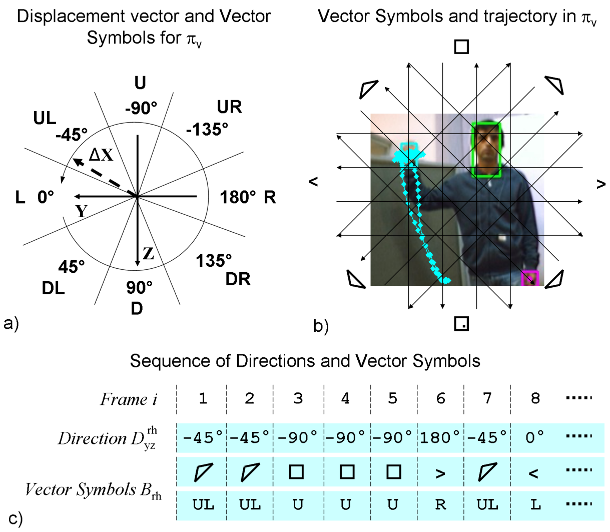


Fig. 2. *Vector Symbols* $B$ of $\pi_v$ and $rh$ for *byebye*. a) The *displacement vector* $\Delta\mathbf{X}$ is converted into the *Vector Symbol* $B_{rh}$. b) Grid of *Vector Symbols* superimposed on the movement trajectory. c) The continuous computation results in a stream of *Vector Symbols*.

## A. Space

The *Space* component presents the different concepts to describe the pathways of human movements inside a frame of reference, when "carving shapes in space" [7]. *Space* specifies different entities to express movements in a frame of reference determined by the body of the actor. Thus, all of the presented measures are relative to the anthropometry of the actor. The concepts differ in the complexity of expressiveness and dimensionality but are all of them reproducible in the 3-D Cartesian system. The following definition is taken from Choreutics (see [10]) and will help to contextualize the first 2 stages of experimental results. Choreutics defines the *Three Planes* - *Door Plane* (vertical) $\pi_v$, *Table plane* (horizontal) $\pi_h$, and the *Wheel Plane* (sagittal) $\pi_s$. We have based our approach on the idea of mapping 3-D trajectories to 2-D planes, though this work does not stick with the 'static' concept of principal planes.

The direction symbols used in the 'original' Labanotation encode a position-based concept of space. As absolute positions demand a higher accuracy from the sensory data we have adopted an earlier concept of Laban. Recently, Longstaff [11] has translated a concept of Laban which is based on lines of motion rather than points in space into modern-day Labanotation. Longstaff coined the expression *Vector Symbols* to emphasize that they are not attached to a certain point in space.

Figure 2 b) shows the *Vector Symbols* presented in [11] as a grid superimposed on a movement trajectory. As shown in Fig. 2 b) we use a coarse (45 degree) discretization of the *displacement vector* $\Delta\mathbf{X}$. For the ease of communication and implementation we use letters to represent the *Vector Symbols*. The correspondence between the *direction* $D$, the signs taken from [11] and the letters can be seen in Fig. 2 c). The figure shows the stream of *Vector Symbols* generated from a *byebye* movement of right hand $rh$ observed in the *Door Plane* $\pi_v$.

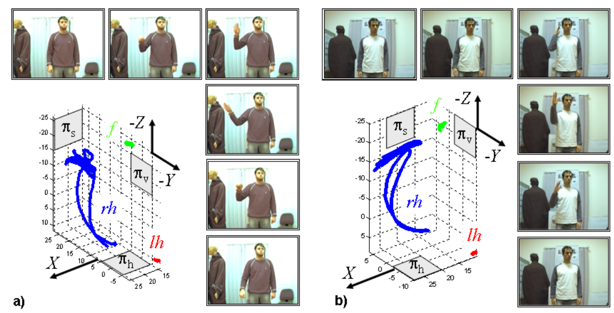


Fig. 3. Recorded 3-D trajectory and observed image sequence for a)a horizontal waving (*byebye*) and b) a sagittal waving (*nthrow*).

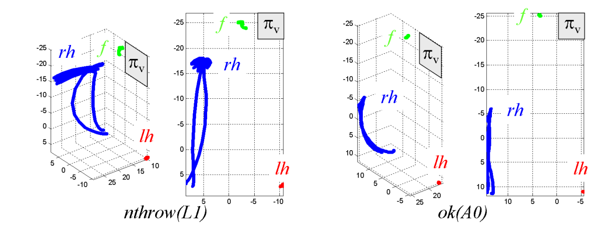


Fig. 4. Two movements with a potential confusion when observed from the *Door Plane* $\pi_v$ alone. The *nthrow* movement looses his oscillatory character from a frontal view and can hardly be distinguished from lifting the hand to make the *ok* gesture. The cipher in brackets indicate person and trial

## B. Database of Expressive Movements

To test our approach we have chosen a set of movements from our database. Some of the movements are based on suggestions mentioned in [7] and [6]. As shown in the example of Fig. 3 some movements can only distinguished if the orientation of the observer is known. In the example the *byebye* gesture represents a horizontal waving, while *nthrow* represents a sagittal waving. In the case of *byebye* the signal can be described primarily by a sequence of $R$ and $L$ *Vector Symbols*, while *nthrow* would be described primarily by $F$ and $B$ *Vector Symbols*.

When observed from the *Door Plane* $\pi_v$ alone, both movements can still be distinguished, though *nthrow* would produce primarily non-movement (0) *Vector Symbols*. When comparing the movement *nthrow* with *ok* a potential confusion occurs when observing only from the *Door Plane* $\pi_v$ as can be verified in Fig. 4.

Also in the case of the movement *maestro* and *byebye* a potential confusion occurs as they appear quite similar when observed from the *Door Plane* $\pi_v$ see (Fig. 5). Though the plane $\pi_v$ appears as the optimal perspective a view from the side ($\pi_s$) may help to resolve the confusion.

It can be seen, that by using the proposed *Vector Symbol* descriptors an accurate camera calibration is not really necessary. Specially the effect of lens distortions can been

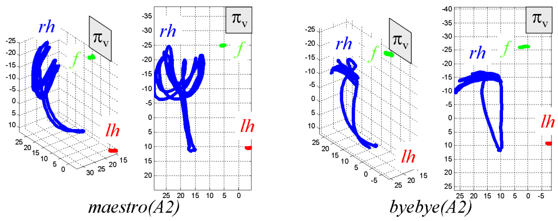


Fig. 5. Two movements with a potential confusion when observed from the *Door Plane* $\pi_v$ alone. The *byebye* as well as the *maestro* movement oscillate from left to right. An additional *Wheel Plane* $\pi_s$ observation might help to distinguish. The cipher in brackets indicate person and trial

neglected for the following reasons: i) The *Vector Symbols* are not depending on absolute positions nor on the magnitude of the displacement and ii) the *Vector Symbols* are discretized with a coarse resolution of 45 degree for a 360 degree range.

## IV. Probabilistic model

### A. LMA Space Model

It was shown in Fig. 2 that the angular values of the *directions* $D$ are translated into the *Vector Symbols* $B_{bp}$ of the *Door Plane* $\pi_v$. The index $bp$ corresponds to the $bodypart$ like the right hand $rh$, the left hand $lh$ and the head (face) $f$. The *Vectors Symbols* receive one additional value, i.e. the indication of no movement $v = 0$. As we describe the spatial pathway of a movement by 'atomic' displacements, we refer to the *Vectors Symbols* sometimes as *atoms*. Movements which are parallel to one of the axes are expressed as up, down, left and right movement resulting in the values $U$, $D$, $L$ and $R$. This represents the concept of *Pure Dimensional Movements* within LMA, while the concepts of *Pure Diagonal Movements* and *Deflections* are described as combinations of *Pure Dimensional Movements*. The temporal dependency of the *Vectors Symbols* is indicated by the frame $I$. The variables and their sample space are shown in (1).

$$\begin{aligned} M \in &\{maestro, \ldots, nthrow\} \; \langle 8 \rangle \\ I \in &\{1, \ldots, I_{max}\} \; \langle I_{max} \rangle \\ B_{bp} \in &\{O, U, UR, R, DR, D, DL, L, UL\} \; \langle 9 \rangle \end{aligned} \tag{1}$$

The *Space model* assumes that each movement $M = m$ produces a certain *atom* $B_{bp} = b$ at a certain point in time, i.e. frame $I = i$ and for a certain $Bodypart$ $bp$. This dependency is also reflected in the Bayesian-net of Fig. IV-A. In this model a certain movement $m$ is 'causing' the *atoms* $a$, $b$ and $c$ at the frame $i$. The *evidences* that can be measured are the *atoms* $b$ and the frame $i$. The movement $M$ is associated with the *concept space*, while the *Vector Symbols* are part of both, the *Laban space* and the *physical space*. The frame $I$ is given by the system as some kind of clock and thus regarded as a 'pure' low-level feature from the *physical space*. The model might be applied to any number of body parts $bp$ which are treated as independent evidences

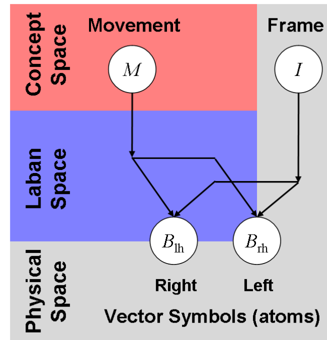


Fig. 6. Bayes-Net for the *Space* component of LMA. The movement $M$ belongs to the *concept space* while the *Vector Symbols* are part of both, the *Laban space* and the *physical space*. Their instances are in the left and right hand. The frame $I$ is associated with the *physical space* only.

and thus expressed through a product as shown in the joint distribution of (2).

$$\begin{aligned} &P(M\ I\ A\ B\ C) \\ &= P(M) \quad P(I) \quad \prod_{bp} \{P(B_{bp} \mid M\ I)\} \end{aligned} \tag{2}$$

The joint distribution contains several distributions. All distributions belong to one of the following groups: i) Distributions that can be determined by 'expert wisdom' and ii) distributions that need to be 'learned'. In the first group priors like $P(M)$ and $P(I)$ can be found. Through 'expert wisdom' we state that all movements are equally likely to occur and thus a uniform distribution is assigned to $P(M)$. In the second group the distribution $P(B \mid M\ I)$ can be found. In our case 'learning' means that trials with a known label are fed into the system which in return identifies the parameters of a chosen distribution.

### B. Continuous classification of movements

Continuous update of the believe is a desirable characteristic of Human-Machine Interaction. With this the system can continuously refine his classification results through the newly incoming evidences. The previous step of learning provided us with the possibility to determine the probability that the *atom* $B$ has value $b$ given a frame $i$ from all possible frames $I$ and a given a movement $m$ from all possible movements $M$, i.e. $P(b \mid m\ i)$. The table $\mathbf{P}(B \mid M\ I)$ holds the probability distribution for all possible values of *atom* $B$ given all possible movements $M$ and frames $I$.

Knowing the conditional probability $\mathbf{P}(B \mid M\ I)$ together with the prior probabilities for the movements $\mathbf{P}(M)$ we are able to apply Bayes rule and compute the probability distribution for the movements $M$ given the frame $I$ and the *atom* $B$ with

$$P(M \mid I\ B) \propto P(M)P(B \mid M\ I) \tag{3}$$

It is possible to compute how likely it is that an observed sequence of $n$ *atoms* was caused by a certain movement $m$. An example for this stream of *atoms* was shown in 2 for the movement $m = byebye$. To compute the *likelihood* we assume that the observed *atoms* are independently and identically distributed (i.i.d.). In (4) the sequence of $n$ observed values for atom $b$ is represented by $b_{1:n}$. For each movement $m$ the joint probability will be the product of the probabilities from frame $i = 1$ to $i = n$, where the $j$th frame of the sequence is indicated by $i_j$.

$$P(b_{1:n} \mid m\ i_{1:n}) = \prod_{j=1}^{n} P(b_j \mid m\ i_j) \tag{4}$$

We can formulate (4) in a recursive way and for all movements $M$ and get

$$P(b_{n+1} \mid M\ i_{1:n+1}) = P(b_n \mid M\ i_{1:n})\ P(b_{n+1} \mid M\ i_{n+1}) \tag{5}$$

The *likelihood computation* (5) can be plugged in our question (3). Assuming that each frame $i$ a new observed direction symbol arrives we can continuously (online) update our classification result.

$$P(M_{n+1} \mid i_{1:n+1}\ b_{1:n+1}) \propto P(M_n)\ P(b_{n+1} \mid M\ i_{n+1}) \tag{6}$$

We can see that the prior of step $n+1$ is the result of the classification of step $n$. Given a sufficient number of evidences (atoms) and assuming that the learned tables represent the phenomenon sufficiently good, the classification will converge to the correct hypothesis. The final classification result is given by the *maximum a posteriori* (MAP) method.

## V. Implementation

### A. Human Movement Tracking

For active sensing we use a 6-DoF magnetic tracker to provide 3-D position data with a sufficiently high accuracy and speed (50Hz). The active sensors are attached to specific body parts, which can also be detected by the visual tracker, i.e. left hand $lh$, right hand $rh$ and face $f$.

The visual tracker performs skin-color detection and object tracking based on the continuously adaptive mean shift (CAMshift) algorithm presented in [3]. CAMshift extends the mean shift functionality by being adaptive to the position and size of a color object. Based on a learned Histogram for skin color a probability will be associated to each pixel that it belongs to a skin colored object. The position of a skin colored object is found by association with the mean of a distribution inside a search window.

It can be seen, that the visual tracker yields a 2-D position that represents the hands not always accurately. The process produces the mean of a 2-D distribution of pixels that have skin color. This yields the center of objects like the hands only under perfect conditions. Changes of the lighting conditions and the shape can shift this 'center' to any position on the object.

### B. Results

The experimental database consists of a set of 8 movements, performed by 4 persons with an average of 4 trials per person for each of the movements, which results in approximately a 160 trials. Two cameras and a commercial motion capture device were used to acquire the data. The 3-D movement data is mapped to planes in 2-D according to the orientation of the cameras. For the Classification, only camera data is used. The following three experiments have been conducted.

TABLE I

CONFUSION TABLE FOR THE SET OF MOVEMENTS WHEN OBSERVING ONLY FROM THE *Door Plane* $\pi_v$.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | $\sum_{errors}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 Lunging | 7 | | | 5 | | | | 1 | 6 |
| 2 Maestro | | 5 | | | | 8 | | | 8 |
| 3 Stretch | | | 12 | | | | 1 | | 1 |
| 4 Ok | | | | 7 | 1 | | 5 | | 6 |
| 5 Point | | | | 1 | 10 | | 1 | | 2 |
| 6 Byebye | | | | | | 13 | | | 0 |
| 7 Shake | | | | 4 | | | 9 | | 4 |
| 8 Nthrow | | | | 4 | | | 1 | | 5 |
| | | | | | | | | | 32(of 95) |

TABLE II

CONFUSION TABLE FOR THE SET OF MOVEMENTS THAT WERE LEARNED IN THE *Door Plane* $\pi_v$ BUT OBSERVED FROM A 45 DEGREE PERSPECTIVE.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | $\sum_{errors}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 Lunging | 10 | | | 1 | 2 | | | | 3 |
| 2 Maestro | | 1 | | | | 12 | | | 12 |
| 3 Stretch | | | 7 | 2 | | | 4 | | 6 |
| 4 Ok | | | | 1 | 10 | | 2 | | 12 |
| 5 Point | | | | 2 | 10 | | | | 2 |
| 6 Byebye | | | | | | 13 | | | 0 |
| 7 Shake | | | | 3 | 2 | 2 | 6 | | 7 |
| 8 Nthrow | | | | | | | 5 | | 5 |
| | | | | | | | | | 47(of 95) |

In the first experiment a single camera was mounted in position parallel to the *Door Plane* $\pi_v$. After computing the projective matrix, a 2-D projection of the movement data was retrieved and learned for this specific orientation. During the Recognition step movements were performed in front of the camera, yielding the results shown in Table V-B. The results show the typical confusions that occur when observing in the *Door Plane* $\pi_v$ alone. The confusion between *maestro* and *byebye* can be verified in Fig. 5 and the the confusion of *nthrow* and *ok* in Fig. 4. With a total of 32 wrong classifications, the classification rate is 66%.

In the second experiment the influence of a distorted perspective was tested. The same single camera was rotated by 45 degree around the performer but now new perspective was learned. During the Recognition step movements were performed yielding the results shown in Table V-B. As can be expected most of the results got worse yielding a classification rate of only 51%.

For the third experiment a second camera was mounted in a position parallel to the *Wheel Plane* $\pi_s$. After computing the additional projective matrix, the 2-D projections of the movement data were retrieved and learned for these specific orientations. Additional cameras and perspectives might be added accordingly. During the Recognition step movements

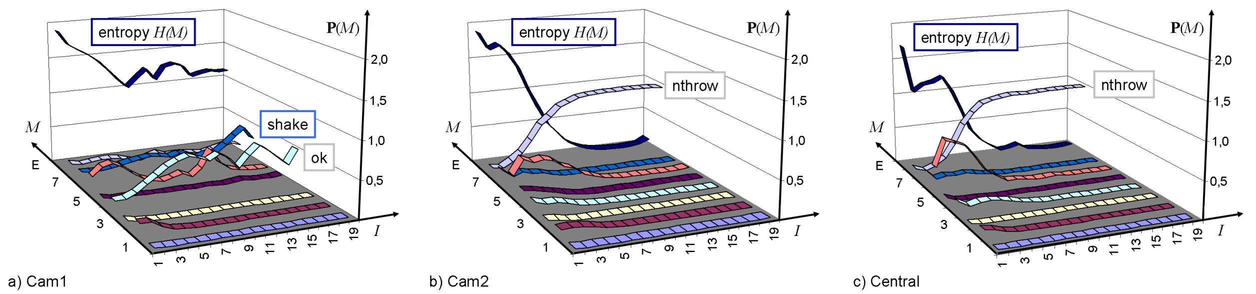


Fig. 7. Continuous update of the probability distribution for the movements while performing *nthrow* a) $Cam1$ aligned with $\pi_v$ produces a high entropy. b) $Cam2$ aligned with $\pi_s$ classifies the movement correctly. c) The fusion follows the belief of $Cam2$ due to its higher certainty.

TABLE III

CONFUSION TABLE FOR THE SET OF MOVEMENTS WHEN OBSERVING FROM THE *Door* $\pi_v$ AND *Wheel Plane* $\pi_s$.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | $\sum_{errors}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 Lunging | 11 | | 1 | 1 | | | | | 2 |
| 2 Maestro | | 11 | | | | 2 | | | 2 |
| 3 Stretch | | | 12 | | 1 | | | | 1 |
| 4 Ok | | | | 7 | 2 | | 4 | | 6 |
| 5 Point | | | | 1 | 11 | | 1 | | 2 |
| 6 Byebye | | | | | | 13 | | | 0 |
| 7 Shake | | | | 3 | 1 | 1 | 8 | | 5 |
| 8 Nthrow | | | | | | | | 5 | 0 |
| | | | | | | | | | 18(of 95) |

were performed in front of camera one and sideways to camera two. The two independent results were then fused in a 'central unit', yielding the results shown in Table V-B.

Concerning the classification rate the results improved to 81%.

This system yields and important characteristic for human-machine interaction. In [2] the characteristic of anticipation for human-robot interaction scenarios was introduced and three factors of anticipation were defined. Similar to their measure of *anticipation-confidence* we compute the entropy $H(M)$ of the movement variable $M$ as a measure of 'certainty'. With this, for each observer system (Cam1 and Cam2) the certainty in their belief can be calculated. Figure 7 shows the evolution of probabilities and the entropy for the observer and central systems during the performance of a *nthrow* movement.

It can be seen that the frontal observer (*Cam1*) is quite uncertain in his hypothesis (*ok* and *shake*) producing an entropy bigger than 1.0. The sideway observer (*Cam2*) which has a clear clearer view of the performance has a entropy converging towards 0. I this case the central system (Central) will mainly follow the belief of *Cam2*.

## VI. CONCLUSIONS AND FUTURE WORKS

The results show that the concept of movement classification based on LMA can be applied to multi-ocular systems. The concept of first recording labeled 3-D movement data using a active sensor and then classifying an unknown movement using computer vision performed well. Using the concepts of *Vector Symbol* descriptors and a Bayesian approach allowed a less accurate calibration resulting in a higher robustness. The Bayesian fusion of multio-cular system allows the introduction of certainty using the measure of entropy. Problems that occur during visual tracking, e.g. occlusions, can be addressed by taking into account the certainty of an observation.

We are currently implementing the Bayesian models for the *Effort* and *Shape* component of the LMA. With a growing database (HID) we can evaluated classification and anticipation of expressive movements. Once evaluated, we want to put our attention to manipulatory movements and the use of LMA as a cue to describe objects properties. A parallel path follows the goal to improve visual tracking by high level knowledge derived from the LMA *Space* component.

## REFERENCES


[1] T. B. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," *CVIU*, vol. 81, no. 3, pp. 231–268, 2001.

[2] J. Rett and J. Dias, "Human-robot interface with anticipatory characteristics based on laban movement analysis and bayesian models," in *Proceedings of the 2007 IEEE 10th International Conference on Rehabilitation Robotics*, ser. Noordwijk, The Netherlands, June 2007.

[3] G. R. Bradski, "Computer vision face tracking for use in a perceptual user interface," *Intel Technology Journal*, no. Q2, p. 15, 1998.

[4] R. Rosales and S. Sclaroff, "Learning and synthesizing human body motion and posture," in *Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, 2000, pp. 506–511.

[5] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[6] L. Zhao, "Synthesis and acquisition of laban movement analysis qualitative parameters for communicative gestures," Ph.D. dissertation, University of Pennsylvania, 2002.

[7] I. Bartenieff and D. Lewis, *Body Movement: Coping with the Environment*. New York: Gordon and Breach Science, 1980.

[8] L. Zhao and N. I. Badler, "Acquiring and validating motion qualities from live limb gestures," *Graphical Models*, vol. 67, no. 1, pp. 1–16, January 2005.

[9] A. Foroud and I. Q. Whishaw, "Changes in the kinematic structure and non-kinematic features of movements during skilled reaching after stroke: A laban movement analysis in two case studies," *Journal of Neuroscience Methods*, vol. 158, pp. 137–149, 2006.

[10] R. Laban, *Choreutics*. London: MacDonald & Evans., 1966.

[11] J. S. Longstaff, "Translating vector symbols from laban's (1926) choreographie," in *26. Biennial Conference of the International Council of Kinetography Laban, ICKL, Ohio, USA*, 2001, pp. 70–86.